Scanners

ICT Training

June 2001

Job Aid

Prepared by Jane Costello

Contents

How scanning works	1
Points to keep in mind	2
Basics of scanning	3
Scanning text—OCR	4
Using Canon scanner	5
OmniPage	6
Scanning text	8
Canoscan	8
Glossary	9
Activities	10
Trainer Notes	11

How scanning works

Scanning is a process of sampling and recording the dots and colour values that make up an image. The image is scanned and sampled line-by-line, and the results are described as a bitmap file. Images are typically scanned at resolutions between 50 to 1200 Dots Per Inch (dpi). Bitmap files are commonly used in paint programs or image processing programs like PhotoShop or Appleworks.

The most popular type of desktop scanner is the flatbed scanner, so called because of its flat, glass platen (or bed) which serves as both the scanning area and surface for laying objects down to be scanned.

What can be scanned?

Two-dimensional opaque objects up to the size of a flatbed scanner (8 1/2 x 14) can be scanned in one pass. Materials to be scanned include:

- photos
- artwork
- drawings
- text
- transparencies
- newspaper clippings

Note: Transparent materials, such as slides or negatives, can be scanned with a special slide scanner or a flatbed scanner that has a transparency adapter.

CSLP

Points to keep in mind before beginning scanning

- Make sure the scanner glass is free of dust, fingerprints, and smudges. Clean it with a soft, lint-free cloth, and a mild, non-alcohol based cleanser.
- Determine if the artwork is either a photograph or a print. If you can see little dots of colour making up the image, then it is a halftone print.
- Consider how the image will be used.
- Some older scanners work better if they are allowed to 'warm up' before scanning.
- Some scanners may have a 'hot spot' where light is more intense. Learn where this hot spot is and avoid it.
- When scanning for screen display you need only be concerned with pixel dimensions. A computer monitor displays approximately 96 pixels per inch. So, if you want an image to appear approximately 1 1/2 by 2 inches on a web page, scan it at 150 by 200 pixels.
- Use no more than 100 dpi/ppi when scanning for screen display.
- Measure dimensions in pixels. 100 pixels on screen is approximately one inch.

Basics of scanning

Please use the user's manual of the specific equipment you are using to locate the specifications for the controls of your scanner.

The following steps are typically followed when scanning:

1. Choose the material to scan.
2. Orient your original on the glass properly. Scanners differ about placement, but usually have some indicator in the corner where you need to position the original.
3. Locate the scanner's sweet spot. A scanner has a sweet spot that can provide sharper focus and, therefore, better detail and truer colour. To test for sweet spots, select a high-quality image with good contrast, fine detail, and bright colours. Scan it several times, with the image on a different part of the platen each time. Open all the scans in your image editor and arrange the windows so you can simultaneously view the same section of each scan.
4. Set the program to scan in black-and-white, greyscale, half-tone or colour, depending on your needs.

5. Set the 'resolution' for the final output in dpi (dots per inch). If the final use of the scan will be for viewing on a monitor, 72 dpi is sufficient. For printing, you will want to scan with a higher dpi, but remember that the file will be significantly larger.
6. Set the 'magnification' for the scan. It is usually indicated with a percentage. 100% is the same size as the original, 50% is half the size of the original, 200% is twice the size of the original, etc..
7. Set the 'exposure' for the scan. Make adjustments in the brightness and contrast of the scan.
8. 'Prescan' the original. A miniature view of whatever is on the glass will appear on screen. You may want to prescan the image before you make any adjustments.
9. 'Select' the area you want to actually scan on the prescan image. This is important because the prescan miniature will show everything and you might only want a portion. To do so, drag the handles of the broken line that appears around your image to include only the final scan area.
10. 'Scan' the image. Usually there is a button.
11. If the program you are scanning with allows it, you may choose to make 'adjustments' in various exposure, contrast, and colour controls at this time.
12. After you scan an image, save it as a TIFF or PSD file. You can convert it afterwards. Do not throw out your original scan after the conversion. If you make changes later, perform the edits in the original and then convert to JPEG. Resaving JPEG files removes even more data, thereby keeping the file size small.

Points to keep in mind when scanning in colour:
An image has three basic colour categories: highlights, shadows, and midtones. If the portrait has a green cast, correct midtones first. Once you have removed the greens from the midtones, you would work on the highlights and then on the shadows. Be sure to check the balance of all three categories–adjusting one range of colours sometimes affects the others.

Scanning text: What is OCR?

Optical Character Recognition (OCR) is a process of scanning printed pages as images on a flatbed scanner and then using OCR software to recognize the letters as ASCII text. The OCR software has tools for both acquiring the image from a scanner and recognizing the text.

Ideal source material for OCR

OCR works best with originals or very clear copies and mono-spaced fonts like Courier. If you have choices, use the following source material:

- 12 point or greater font size.
- Black text on a white background.
- A clean copy; not a fuzzy multi-generation copy from a copy machine.
- Standard type font (Times, New Roman, etc.) Fancy fonts may not be recognized.
- Single column layout.

OCR limitations

- Using text from a source with font size less than 12 points or from a fuzzy copy will result in more errors.
- Except for tab stops and paragraphs marks, MOST document formatting is lost during text scanning, (Bold, Italic & Underline are sometimes recognized).
- The output from a finished text scan will be a single column editable text file. This text file will always require spell checking and proofreading as well as reformatting to desired final layout.
- Scanning plain text files or printouts from a spreadsheet usually works, but the text must be imported into a spreadsheet and reformatted to match the original.

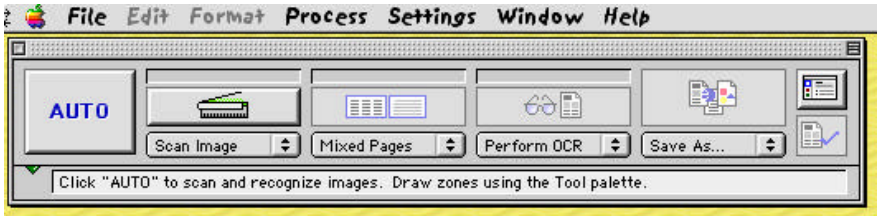
What source material does not work well for OCR?

- Forms (especially with boxes and check boxes)
- Very small text
- Multi-generation fuzzy or blurry copies from a copy machine
- Mathematical formulas
- Draft copies of documents with hand-written revisions
- Fancy text and unusual fonts
- Handwritten text

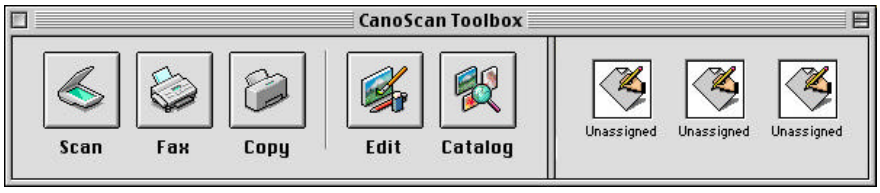
Using the Canon CanoScan N1220U scanner

Note: The following directions are for use specifically with the Canon CanoScan N1220U scanner that is found in this lab. While the process is basically the same for all flatbed scanners, some of the details will not match. You need to refer to the user's manual of the specific equipment you are using in your own lab to locate the specifications for the controls of your scanner.

There are two software applications to use for scanning with the CanoScan N1220U, Omnipage

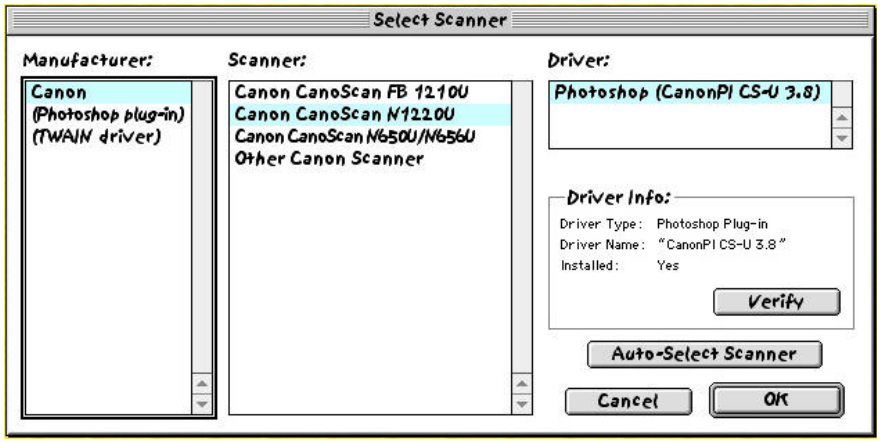

File Edit Format Process Settings Window Help
AUTO
Scan Image
Mixed Pages
Perform OCR
Save As...
Click "AUTO" to scan and recognize images. Draw zones using the Tool palette.

and CanoScan.

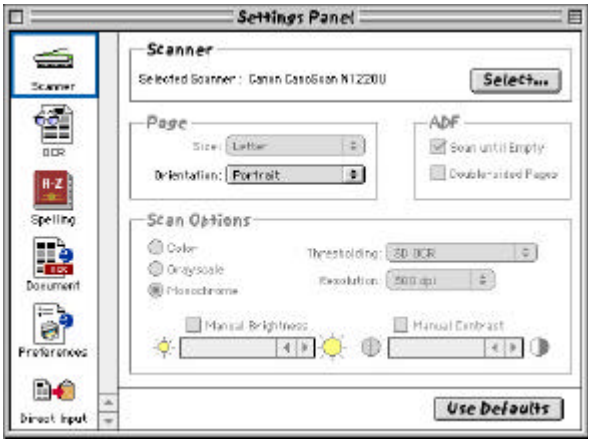


OmniPage

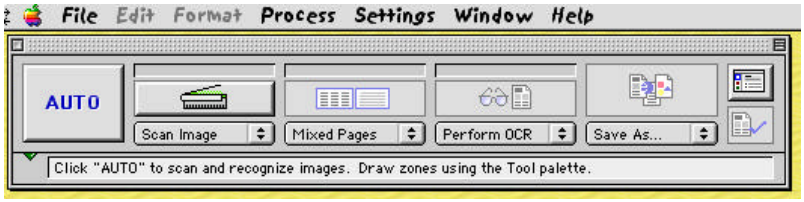
1. Open the OmniPage software.
2. Verify that the scanner has been selected by going to 'Settings', 'Select Scanner'.
   The following dialog will appear:



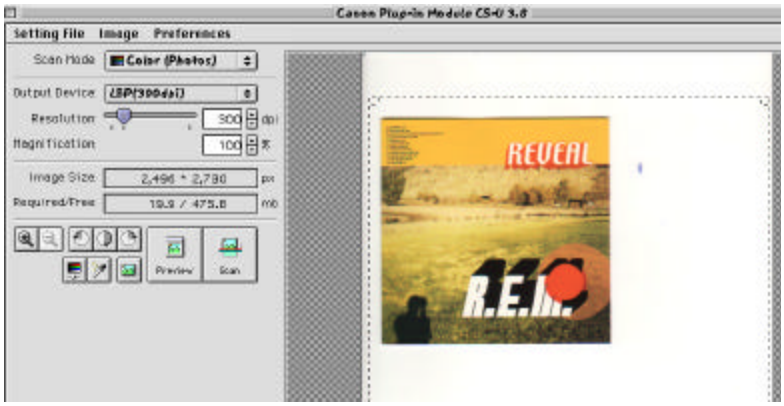
3. Select the appropriate settings through 'Settings Panels' on the menu bar. Verify the settings for the 'Scanner', 'OCR', 'Spelling, Document', 'Preferences' and 'Direct Input' features. Close the dialog by clicking on the top left-hand corner.
4. To scan an image, place the item to be scanned on the glass. Be sure to place it in the correct position. Click on the 'Scan Image' button.

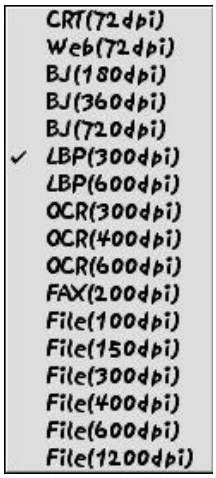


The scanner environment will appear.


File Edit Format Process Settings Window Help
AUTO
Scan Image
Mixed Pages
Perform OCR
Save As...
Click "AUTO" to scan and recognize images. Draw zones using the Tool palette.

5. Click on 'Preview' to preview the image you are scanning.

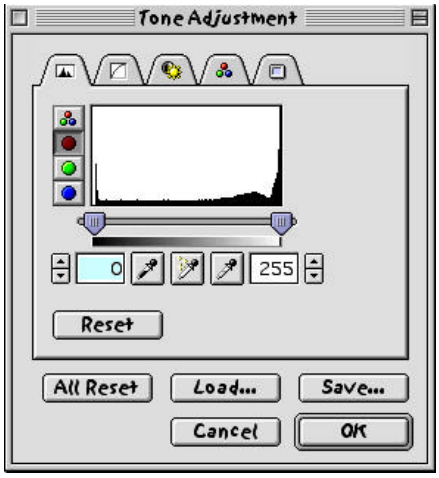




6. Choose the scan mode from the list.

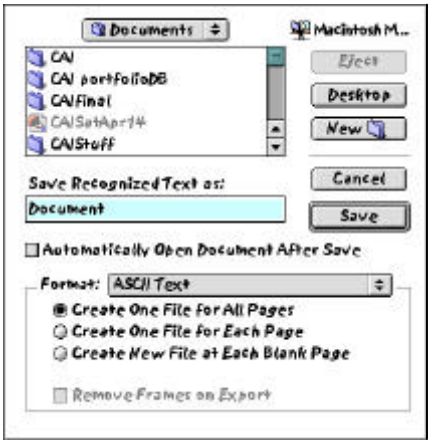
7. Choose the output device from the list. (Note: for most images 300 dpi is suitable.)

8. Make any adjustments to the resolution and magnification. Use the tone adjustment tool to adjust the colour settings.



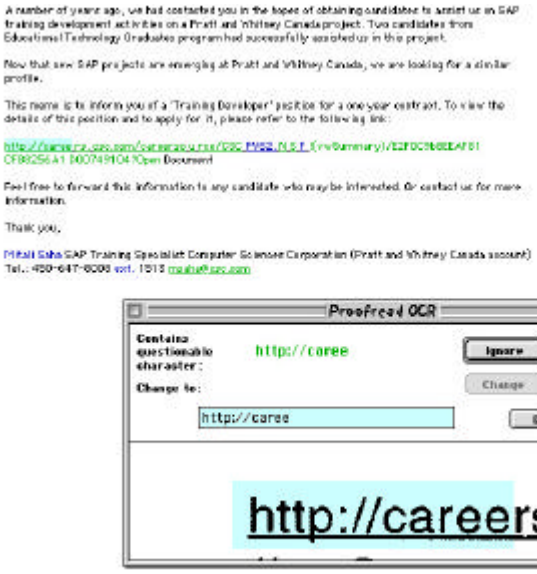
9. Click on the 'Scan' button to make the final scan. The scanned image (with all applied settings) appears on the screen.

10. Save the image in the desired location by choosing 'File', 'Save as'. The Save dialog will appear.



Scanning text

1. This is basically the same procedure as for scanning an image. However, choose black and white scan mode.
2. Once the scanned image appears on the screen, choose 'Perform OCR' from the toolbar. Omnipage now knows that the image is text and will attempt to section it off and translate the contents of the scanned item into a text file.
3. Verify the spelling of the text by clicking on the 'Proofread' button on the toolbar. The following dialog appears:



4. Once the proofreading is complete. Save the file.

You can also copy and paste this text file into a word processor and save it there.

CanoScan

1. Open the CanoScan Toolbox software from your computer desktop by pressing on the scan button on the front of the scanner.
2. Click on the 'Scan' button on the CanoScan Toolbox. The scanner scans the image and displays the image in the CanoScan editing environment. This environment contains three tool pallets. They are similar to those found in a basic paint program.
3. Edit your scanned image as required and save the image to your folder.

Glossary

Halftone printing: A method for reproducing colour and black and white graphics using dots of ink. The most obvious examples are the colour photographs in newspapers.

Optical Character Recognition (OCR): is a process of scanning printed pages as images on a flatbed scanner and then using OCR software to recognize the letters as ASCII text. The OCR software has tools for both acquiring the image from a scanner and recognizing the text.

Optical resolution: This is the actual, physical ability of the scanner to discern individual dots. A 300 dpi scanner can accurately reproduce a 300 dots per inch image.

Scanning: A process of sampling and recording the dots and colour values that make up an image. The image is scanned and sampled line-by-line, and the results are described as a bitmap file.

Sweet spot: A scanner has a sweet spot that can provide sharper focus and, therefore, better detail and truer colour. To test for sweet spots, select a high-quality image with good contrast, fine detail, and bright colours.

Transparency adapter: This is an optional lid for flatbed scanners that has a light source inside, usually florescent. It is useful for scanning transparencies, x-rays, etc. It can be used for scanning slides as well, but a slide scanner will do a vastly better job.

Twain support: Twain is a software interface standard developed by Hewlett-Packard for interfacing scanners to productivity software. The general idea is that a device with a twain driver can be used with any software that supports the Twain standard. Twain support is mainly used in the Windows environment.

Scanners Activities

Using the directions previously outlined in this job aid, complete the following activities. Save them to your account. You may wish to use some of these items in the Web page Building workshop.

Activity 1
Scan a photograph.

Activity 2
Scan a drawing.

Activity 3
Scan a newspaper clipping.

Activity 4
Scan a page of text.

Activity 5
Scan an object.

Scanners

Trainer Notes

Helpful scanning tips

- The quality of your original will directly affect the quality of the digital image you can create, so choose your originals carefully.
- Think ahead to what the end use of the image will be.
- Most crashes with graphics programs are due to insufficient memory.
- If you are printing the image, laser printers do a terrific job with 300 dpi, but when inserting a file in a word processing or presentation document, look for the option to link or publish the graphic file instead of pasting the whole file, especially since 300 dpi graphic files are large Threshold Specifies how many numbers of samples in an image will be sharpened. A setting of 0 will affect every sample whereas a setting of 50 will affect almost none of the samples. Highly detailed images, such as line art, require a setting of 3 whereas portraits look best with a setting between 5 to 9.
- To clean your scanner's platen, a quick once-over with a soft cloth, paper towel, or sheet of newspaper might be all it takes to remove tiny bits of dust that can be visible in a high-resolution scan. For fingerprints, any type of glass cleaner will work, but check the scanner manual first to be sure that the platen is cleanser-safe.

CSLP